An Effective Approach to an Image Retrieval using SVM Classifier

Mohd. Aquib Ansari¹*, Manish Dixit², Diksha Kurchaniya³, Punit Kumar Johari⁴

¹Dept. of CSE/IT, Madhav Institute of Technology and Science, Gwalior, India  
²Dept. of CSE/IT, Madhav Institute of Technology and Science, Gwalior, India  
³Dept. of CSE/IT, Madhav Institute of Technology and Science, Gwalior, India  
⁴Dept. of CSE/IT, Madhav Institute of Technology and Science, Gwalior, India

*Corresponding Author: aaquib.mits@gmail.com, Tel.: +91-8269786126
Available online at: www.ijcseonline.org

Abstract— Content Based Image Retrieval (CBIR) is an important and challenging research area of digital image processing for browsing and retrieving the images from the wide range of image database. The basic requirement of the CBIR system is to retrieve the relevant information from large image database following to query image with higher system performance. In this paper the two step strategy is developed in which the first step is to extract the low level or pixel level features of the image by using color, texture and shape descriptor. Here an efficient fused feature extraction method based on color and edge directivity descriptor (CEDD) for extracting the color as well as texture features and two level discrete wavelet transform (2D-DWT) for extracting the shape feature of the image is proposed. While, in the second step the SVM classifier is used to classify the images into different classes and to handle irrelevant examples. For retrieving the similar images following to query image the Euclidean distance similarity measurement is used. This fused and classified based proposed scheme applied on different image databases and proved that it is providing better results over various existing methods and individual approaches.

Keywords— CBIR, feature extraction, color and edge directive descriptor, 2 level discrete wavelet transform, support vector machine, similarity measurement

I. INTRODUCTION

There is a good saying about picture said by Chinese philosopher “A Picture is worth a thousand words” means that a picture can convey more than words. During the last few years, the digital image technology is improving day by day. The images are using in the various domains like crime prevention, medical, healthcare, web searching, industries etc. So, the growth of the digital images are increasing rapidly. This huge number of digital images has made challenges to the system to retrieve or browse the relevant images from the large digital image repository efficiently. To fulfil this requirement the Content Based Image Retrieval (CBIR) [1] came into frame in 1992.

Content Based Image Retrieval (CBIR) is a system which retrieves the relevant images from the huge image repository automatically following to query image using the visual features. The CBIR mainly consist of two function. The first is feature extraction, in this step the set of features are extracted from the image which can uniquely represent the image. The second is similarity measurement which measures how similar the retrieved images are with respect to query image. Similarity measurement measures the similarity between the query image and images in the image database [1] [2].

The features are extracted by analysing the actual content of image rather than the descriptor or keyword attached to it. CBIR uses the low level features or pixel level features like color, texture and shape features to extract the appropriate information of the image. In CBIR, first the features are extracted for each image of image database and stored in a separate database called the feature database. In second step the user fires the query image as input image. The features are extracted for this query image and stored in a vector called query image feature vector. On the basis of similarity measurement this query image feature vector compares with each feature vector of feature database and retrieved the similar images corresponding to query image as result [2] [3]. In figure 1 the overall basic architecture of CBIR is shown.

Pre-processing [2] is an important step in CBIR which is used before the feature extraction process. In pre-processing step the image can be resized, segmented, filtered (to remove the noises), normalized, converted into required image space etc.
Figure 1. Basic architecture of Content Based Image Retrieval

This paper is organized in various sections. The section II describes the existing work of CBIR. The techniques involved in this proposed approach and the main image retrieval algorithm are described in section III. The section IV shows the experimental results. Section V mentions the conclusion with future scope followed by references.

### II. RELATED WORK

<table>
<thead>
<tr>
<th>S.NO</th>
<th>Year</th>
<th>Author</th>
<th>Proposed work</th>
<th>Database used</th>
<th>Result Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2016</td>
<td>Sawet Somnugpong [5]</td>
<td>Color Correlogram is used to extract the spatial information and Edge Histogram Descriptor is used to extract geometric information of the image.</td>
<td>Wang database used having 1k images with 10 classes.</td>
<td>The proposed system is performing better result in 6 of 10 retrieval class. Need to combine other features with EHD to get better result.</td>
</tr>
<tr>
<td>2</td>
<td>2016</td>
<td>Yogita Mistry [6]</td>
<td>Local binary pattern and Dual Tree Discrete Wavelet Transform are used to extract the features.</td>
<td>Wang database used having 1k images with 10 classes.</td>
<td>Better precision and recall over existing methods. But need color feature to explore more.</td>
</tr>
<tr>
<td>3</td>
<td>2016</td>
<td>Raj Kumar Jain [7]</td>
<td>Hsv color quantization and Prewitt Edge Detection are used to extract color and shape features of image.</td>
<td>Wang database used having 600 images with 6 classes.</td>
<td>This combined algorithm is performing well in precision. But need texture feature to explore the result more.</td>
</tr>
</tbody>
</table>
| 4    | 2016 | Garima Gupta [8] | The color, texture and shape features (Dominant Color Descriptor, Wavelet Transform and Hough Transform) are used to extract essential information of image. Apart these features, for classification RBF-SVM is used. | Facial Images with 444 Images. | Avg. precision=0.69  
Avg. f-measure=0.68  
Avg. accuracy=0.69 |
<p>| 6    | 2016 | Pradnya Vikhar [9] | Edge Histogram Descriptor, Color Auto Correlogram, Color Moments, Gabor Wavelet Transform are combined. Then on this combine approach the | IMAGEVARY database with 1k Images. | Good retrieval results with precision and recall value after applying SVM. |</p>
<table>
<thead>
<tr>
<th>Year</th>
<th>Author(s)</th>
<th>Methodology</th>
<th>Dataset Description</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>Katta Sugamya [10]</td>
<td>SVM Classifier used by color, texture and shape feature extraction for classification.</td>
<td>200 images of zebra and tiger (100 images for each category).</td>
<td>Accuracy and error rate are found to be precise.</td>
</tr>
<tr>
<td>2016</td>
<td>Navneet Kaur [11]</td>
<td>The learning techniques such as KNN, SVM and Bayes classifier are used to classify the image database. Color Moment, Color Histogram and Gabor Wavelet are used to extract color and texture features.</td>
<td>Corel image database with 10k images.</td>
<td>This proposed method is efficient to minimize the semantic gap. This research can be extended by fusion of KNN with any unsupervised algorithm.</td>
</tr>
<tr>
<td>2015</td>
<td>Kommineni Jenni [12]</td>
<td>This work is concentrated on database classification and efficient image representation. Color string coding and string comparison is used to extract the features of the image. SVM classifier is used to classify the image database.</td>
<td>1800 images from Corel photo family.</td>
<td>Proving good result in term of precision and recall. The result may be enhanced by using other optimization techniques.</td>
</tr>
<tr>
<td>2015</td>
<td>Amit Khatami [13]</td>
<td>FCS is used as color feature. The Swarm intelligent based on kmedoids algorithm is used in fire detection field.</td>
<td>Forest Images</td>
<td>This proposed algorithm is fast and efficient to detect fire in the forest.</td>
</tr>
<tr>
<td>2014</td>
<td>Manoj D. Chaudhary [14]</td>
<td>Stationary wavelet transform (SVD) and Edge Histogram Descriptor are used to extract the texture as well as shape features of image.</td>
<td>Wang database used having 1k images with 10 classes.</td>
<td>Average Precision=0.77 but need color feature to explore the result more.</td>
</tr>
<tr>
<td>2013</td>
<td>Swati Agarwal [15]</td>
<td>Here Wavelet transform is used to decompose the image into different components and on these selected components the edge histogram descriptor is used to extract the relevant information of the image.</td>
<td>Wang image database</td>
<td>This combined approach using shape and texture is more efficient than other’s. In the experiment analysis it is also find that Coif4 discrete wavelet function performs better results than other wavelets.</td>
</tr>
<tr>
<td>2012</td>
<td>Raghupati Gali [16]</td>
<td>The Color Histogram, Contourlet and Edge Histogram Descriptor are used as color, texture and shape feature to extract information of image. On the combination of color texture and shape features the Genetic algorithm applied.</td>
<td>Wang database used having 1k images with 10 classes.</td>
<td>Reduced semantic gap.</td>
</tr>
</tbody>
</table>
III. METHODOLOGY

A. Color and Edge Directivity Descriptor
Color and Edge Directivity Descriptor (CEDD) [17] [18] is used to extract the color as well as Edge information of the image. One of the advantage of CEDD is its low computation power in comparison of other MPEG-7 Descriptors. The CEDD contains 2 units for extracting the features of the image like color unit for extracting color features and texture unit for extracting texture feature. HSV color space is used in color unit and YIQ color space is used for texture unit. The set of fuzzy rules are used for fuzzy linking histogram to extract the color feature. The 20 fuzzy rules are applied on 3 input fuzzy system to extract the 10 bins of quantized histogram. For converting the 10 bins of histogram into 24 bins of histogram 4 extra bins are used on 2 input fuzzy system. For the texture unit the five digital filters are used to extract the edge information of the image. These five filters are the part of Edge Histogram Descriptor [19]. For texture unit CEDD has 6 regions and each region has 24 individual regions. Total bins for texture unit includes 24x6=144. Then it combines color and texture information in a histogram. The flowchart of CEDD descriptor is show in figure 3 [18].

B. Discrete Wavelet Transform
Discrete wavelet transform (DWT) [8] [10] [20] is an efficient as well as widely used approach to extract the spatial information of the image. For multi scale image analysis DWT is the best descriptor. Means that if a pattern could not analyse in one resolution it can be analyse in another resolution. DWT decomposes the image into four different sub-bands by passing it low and high pass filter. These sub-bands are approximation band which contains approximated image, horizontal band which contains horizontal details of image, vertical bands which contains vertical details of image and last diagonal band which contains diagonal details of image. From these sub-bands the detailed images measure variations. The approximated image is decomposed again to wavelet sub-bands. Object recognition uses more than one level of decomposition for getting the reduced and meaningful information of image [21] [22]. Here 2 level DWT is used. Feature vector is calculated by analysing the final approximated image.

C. Support Vector Machine
Support Vector Machine (SVM) [8] [9] [10] [11] [21] is supervised learning technique which is used to examine the data and detect the classified pattern used. SVM is a classifier that is used to classify the data into different classes with the help of hyper planes consisting minimum classification error and maximum margin. The main module of SVM is to make hyper planes or group of hyper planes by using the support vectors in a higher dimension space. For nonlinear separable data SVM uses kernel function to classify the images. Here radial basis kernel function is used. Which can classify the large number of data space than linear and polynomial function [21] [23].

D. Similarity Measurement
Similarity measurement is an important step in CBIR to retrieve the similar images following to query image from the large set of image database. Here Euclidean Distance [24] [25] [26] is used as similarity measurement. It is widely used measurement in CBIR because of its efficiency and effectiveness. It compares the feature vector of query image to each feature vector of image database and measures the distance matrix (D) for each image of database. Then distance (D) is sorted in ascending order and first N (number of retrieved image) images are displayed. This distance matrix (D) describes that how much image in the database is similar to query image.

\[
D = \sqrt{\sum_{i=1}^{l}(p_i - q_i)^2}
\]  

Where D is the distance matrix, p and q are feature vector of query image and feature vector of image in the database. l is length of the feature vector.

E. Proposed Algorithm
1. Select the directory of image.
2. Take input image.
3. Pre-process the image by resizing it into 256x284.
4. Apply CEDD feature to extract color and edge information of the image (f1).
5. Apply 2D-DWT to extract spatial information of the image (f2).
6. Combine the both features into single unit f = [f1 f2].
7. Repeat step 2 to step 6 until the last remaining image of image database. After extracting features for image
database store these features in a separate database called feature database.

8. Take query image.
9. Repeat step 3 to step 6 to extract the features from query image and named query image feature vector.
10. Apply Support Vector Machine Classifier to classify the image database into different classes.
11. Apply Euclidean Distance Measurement to display the retrieved images.
12. Calculate Precision, Recall and Accuracy.

F. Proposed Retrieval Scheme

Figure 3. Diagramatic representation of proposed retrieval system

IV. EXPERIMENTAL RESULTS

A. Image Database
This proposed method is employed on 3 different databases such as Wang image database [5] [6] [7] [15], Caltech image database [29] and Corel image database [11] [12]. The experimental analysis is done with 1000 images of Wang image database containing 10 classes, 800 images of Caltech database containing 8 classes and 700 images of Corel database containing 7 classes. The 10 classes of Wang database are African, Beach, Monuments, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains and Food respectively, shown in figure [4]. The 8 classes of Caltech are Ketch, Aeroplane, Faces, Piano, Bikes, Leopard, Bonsai and watch respectively, shown in figure [5]. The 7 classes of animals are taken from the Corel 10k image database. These classes of Corel Database are Bear, Wolf, Lion, Tiger, Dog, Monkey and Cat respectively, shown in figure [6].
Where effectiveness and efficiency describes the high accuracy of the retrieval system and speed of retrieval system respectively. Here precision, recall and accuracy measures are used to evaluate the performance of retrieval system. Precision [11] [28] measures the capability of the system to retrieve the similar images only. Recall [11] [28] measures the capability of system to retrieve all the similar images from the image database.

\[
\text{Precision (P)} = \frac{TP}{(TP + FP)}
\]

(2)

\[
\text{Recall (R)} = \frac{TP}{(TP + FN)}
\]

(3)

The accuracy [10] [11] [28] is evaluated for image classification based system. Accuracy is the ratio of the items classified correctly to all classified items.

\[
\text{Accuracy (A)} = \frac{TP + TN}{(TP + FN + FP + TN)}
\]

(4)

Where, TP (True Positive) = images that are relevant and retrieved.

TN (True Negative) = images that are relevant but not retrieved.

FP (False Positive) = images that are not relevant but retrieved.

FN (False Negative) = images that are not relevant and not retrieved.

C. Result Analysis

This proposed approach using color and edge directivity descriptor, 2 level discrete wavelet transform and support vector machine classifier is implemented on Matlab R2013b on PC core i3 5th Generation 4 GB of Ram capacity. The different databases used for performing experimental analysis of our proposed method are Wang image database, Caltech image database and Corel image database. Here precision, recall and accuracy values are calculated for each category according to equation 2, equation 3 and equation 4 respectively. In the table 2, the 1k images of Wang database are taken with 10 categories of image. The classification is performed for individual approaches as well as combined approach (proposed approach). It finds that this proposed scheme is providing better result over individual classified schemes in term of precision, recall and accuracy.

<table>
<thead>
<tr>
<th>Categories</th>
<th>CEDD + SVM</th>
<th>2D DWT + SVM</th>
<th>Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision</td>
<td>Recall</td>
<td>Accuracy</td>
</tr>
<tr>
<td>African</td>
<td>0.680</td>
<td>0.76</td>
<td>0.822</td>
</tr>
<tr>
<td>Beach</td>
<td>0.638</td>
<td>0.60</td>
<td>0.836</td>
</tr>
<tr>
<td>Monuments</td>
<td>0.740</td>
<td>0.80</td>
<td>0.838</td>
</tr>
<tr>
<td>Buses</td>
<td>0.957</td>
<td>0.90</td>
<td>0.812</td>
</tr>
<tr>
<td>Dinosaurs</td>
<td>0.980</td>
<td>0.98</td>
<td>0.842</td>
</tr>
<tr>
<td>Elephants</td>
<td>0.700</td>
<td>0.70</td>
<td>0.846</td>
</tr>
<tr>
<td>Flowers</td>
<td>0.979</td>
<td>0.94</td>
<td>0.816</td>
</tr>
<tr>
<td>Horses</td>
<td>0.941</td>
<td>0.96</td>
<td>0.822</td>
</tr>
<tr>
<td>Mountains</td>
<td>0.733</td>
<td>0.66</td>
<td>0.816</td>
</tr>
<tr>
<td>Food</td>
<td>0.914</td>
<td>0.86</td>
<td>0.824</td>
</tr>
<tr>
<td>Average</td>
<td>0.826</td>
<td>0.816</td>
<td>0.827</td>
</tr>
</tbody>
</table>

In the table 3, it can be noticed that this proposed method gives better results over other existing retrieval methods, namely Method [28], Method [9] and Method [16].
The experiment is done on Caltech database with 800 images of 8 categories of image. After performing classification the results are shown in table 4. For Caltech database it also proves that this proposed scheme provides better result over individual classified approaches.

In the table 5, the 7 categories of animal images are taken from the Corel database. Each category having 100 images. The Classification is performed to classify the animals into their respective classes. It is shown that our proposed scheme classified the animal images so well in comparative to individual classified approaches.

The comparison graphs between proposed method verses individual methods in terms of precision, recall and accuracy for Wang, Caltech and Corel image database are shown in figure 7, figure 9 and figure 10 respectively. Figure 8 shows the comparison graph between proposed method and existing methods in term precision and recall.
Figure 7. Comparison graph among CEDD, DWT and Proposed method (By using Wang database)

Figure 8. Comparison graph between Proposed Method verses Existing Methods (By using Wang database)

Figure 9. Comparison graph among CEDD, DWT and Proposed Method (By using Caltech database)
Figure 10. Comparison graph among CEDD, DWT and Proposed Method (By using Corel database)

Figure 11. First 20 images retrieved for query image of flower using proposed methodology (By using Wang database)

Figure 12. First 20 images retrieved for query image of ketch using proposed methodology (By using Caltech database)

Figure 13. First 20 images retrieved for query image of bear using proposed methodology (By using Corel database)

V. CONCLUSION AND FUTURE SCOPE

In order to increase the efficiency and effectiveness of the content based image retrieval this paper presented a classification based hybrid approach using CEDD, DWT and SVM classifier. Here CEDD used set of fuzzy rules for fuzzy color histogram to extract color information and also used five digital filters to extract edge information of the image. DWT extracted the spatial information of the image by evaluating the low frequency band of the decomposed image. After extracting features the classification is performed using support vector machine classifier, which classified the images into different classes very efficiently. The Experiments carried out on the different type of image database. On the analysis of outcomes it is found that this proposed approach is performing well in term of precision, recall and accuracy over the individual feature classified approaches and other existing
approaches. This proposed approach touched the average accuracy up to 0.85, 0.90, and 0.78 for Wang, Caltech and Corel image database respectively. For future work other optimization schemes such as ant colony optimization, particle swarm optimization, and combination of both algorithm can also improve the performance of CBIR system more quickly.

REFERENCES


[28] Priyadarshini Patil, and Bhagya Sunag, ”Analysis of Image Retrieval Techniques Based on Content”, International Advance


Authors Profile

Mr. Mohd. Aquib Ansari received the Bachelors degree in Information Technology from SATI, Vidisha, M.P. in 2014. He is currently pursuing his Master of Technology from MITS, Gwalior, M.P. He is an IEEE student member. His research interest include image retrieval. He can be contacted at aqub.mits@gmail.com.

Ms. Diksha Kurchaniya received the Bachelors degree in Computer Science and Engineering from ITM, Gwalior, M.P. in 2014. She is currently pursuing her Master of Technology from MITS, Gwalior, M.P. Her research interest include image retrieval. She can be contacted at diksha5mits@gmail.com.

Mr. Manish Dixit received the B.E. in the computer technology from Barkatullah University, Bhopal and M.E. in Communication Control and Networking from MITS, Gwalior in 1993 and 2006 respectively. He is pursuing Phd from Ragiv Ghandhi Technical University, Bhopal. He is currently working as an Associate Professor in the Department of Computer Science & Information Technology, MITS, Gwalior, India. He has presented various research papers in the National and International Conferences and Journals. He is a Member of CSI, IETE, IAENG. He is a member of IEEE & Chairmen IEEE MP Section. He can be contacted at dixitmits@gmail.com.

Mr Punit Kumar Johari pursed Master of Technology in Information Technology from RGPV, Bhopal in year 2006. He is currently pursuing Ph.D. and working as Assistant Professor in Department of Computer Science and Engineering and Information Technology, MITS, Gwalior since 2010. He is a member of IET & CSI since 2016 & 2015, a life member of the GAMS since 2006 and IAENG since 2012. He has published more than 20 research papers in reputed international journals including Thomson Reuters and conferences including IEEE and it’s also available online. His main research work focuses on Computer Vision, Data Mining and Network Security. He has 12 years of teaching experience and 5 years of Research Experience. He can be contacted at punitbhopal2006@gmail.com.